EAGAN, MCALLISTER ASSOCIATES INC.

NEXGENBUS Initial Model Results

BACKGROUND

The current DoD policy for Acquisition Reform uses Commercial-Off-The-Shelf (COTS) products whenever possible to avoid development costs of new systems acquisition. The Next Generation Instrumentation Bus (NexGenBus) project is currently evaluating commercial standards for a high-speed instrumentation bus standard. These commercial standards need to be tested, evaluated, and possibly changed based on the application. Adapting a commercial standard to make it useable for a specific application (e.g. test instrumentation) requires that the standard be evaluated and tested to insure compliance and to insure that a non-conforming variation to the standard is not created.

PURPOSE

When choosing a standard for NexGenBus there are overlapping areas within the standard and the application. These overlapping areas within NexGenBus contain critical areas that must be tested to verify their compliance with the standard. These critical areas frequently require simulation to evaluate their effectiveness without building a complex lab set-up. These critical areas include new protocols, loading analysis, flow control and error correction. Simulating a protocol or introducing errors in a simulation may be much more cost effective than doing actual testing. This philosophy requires the requisite test areas are done in the lab but the difficult or impossible test points are done in a simulation. This complements the lab testing and saves the cost of expensive test equipment. However, since a simulation is only as good as the parameters that are input, some lab testing is required to provide verification of the simulation.

SYSTEM ANALYSIS AND MODELING OBJECTIVES

The main objective of the NexGenBus initial modeling effort is to simulate the arbitrated loop fiber channel architecture to determine the message throughput and latency. This was done by developing a Visual C++ program called WinBench for the NexGenBus program. The WinBench program uses the driver obtained from the fibre channel PCI card vendor and adds the features needed to verify the initial model of the NexGenBus architecture. The WinBench program has the following features:

- Allows the size of the message to be varied

- Allows the contents of the payload to be defined (ASCII characters)
- Allows a variable delay to be added between messages
- Allows a variable number of messages to be sent
- Allows a continuous message transfer (infinite number of messages)
- Saves the file transferred to use for comparing to original file

NEXGENBUS MODEL REQUIREMENTS DEFINITION

This section pertains only to those requirements that will be demonstrated by simulation. The electrical tests will be completed in a separate effort outlined in the NexGenBus Test Plan for Fibre Channel. The requirements for the initial models are to simulate the arbitrated loop (two nodes only) lab set-up and verify that the simulation can duplicate the throughput and latency demonstrated in the lab. The minimum requirement for data rate on the NexGenBus was determined to be 100 Mbps. However to ensure the bus chosen will meet future requirements it is desirable that the data rate be as high as possible. Fibre Channel's current maximum transmission rate of 1.062 Gbaud yields an unencoded data rate of 800 Mbits per second. The baseline model will be an arbitrated loop topology with two nodes and therefore there will be no measure of synchronicity. The end-to-end latency of the bus transmissions is dependent on the topology, therefore the point-to-point topology will experience the lowest latency. The throughput measurements are calculated using the WinBench program as the time to transmit divided by the bytes received. The lab set-up is used to verify the simulation results.

When the initial model is verified a subsequent model will add capabilities such as different topologies, additional nodes and additional message sources. By using this expanded model, topology changes can be quickly simulated and throughput evaluated. Additional nodes or message sources can easily be added to the model to determine the best placement for the node and its effect on the existing bus. The simulation model is show in Figure 1.

Figure 1 NEXGENBUS Initial Model

COMNET III MODELING TOOL

COMNET III (CACI Products Company, La Jolla CA.) is a network modeling and simulation tool that was used to simulate the NexGenBus data network and estimate the performance characteristics of the NexGenBus lab set-up. The model is created graphically using a Windows NT Graphical User Interface (GUI). Uses of the COMNET software application include:

- Peak loading studies
- Network sizing at the design stage
- Resilience and contingent planning
- Introductions of new hardware or applications
- Evaluating performance improvement options
- Evaluating grade or class of service

Objects representing the various hardware components are created within the application. These objects include communication nodes, switches and links with their parameters edited to match the characteristics of their real world counterparts. The traffic loading and computer workloads are derived from the vendors supplying the hardware. Each node is represented by a processor node icon with identical characteristics to the lab set-up configurations. Connected to each processor node is a network message that originates or responds to the processor node. Name, periodicity, origin, and destination characterize the messages. For the NexGenBus model processor nodes are connected via a fibre channel datalink which is configured to match the characteristics of the lab datalink. Table 1 lists the Configuration Parameters for the model components that were derived from their real-world counterparts.

Table 1: COMNET III Configuration Parameters

Name	Type	RAM (MB)	Buffer Space (MB)	Processor Cycle (usec)
XmitComp	Processor Node	64	1	. 015
RcvComp	Processor Node	64	1	. 015

Name	Type	B.W. (Kbps)	Frame Min.	/ Max.	/ O.H.(bytes)
FibreChannel Link	Pt-to-Pt Link	1062.5	32	2016	32

Name	Type	IAT (msec.)	Command Sequence	ACK. (bytes)
Xmit_Source	App.	. 00075	Local (Trpt) Set_Up_Frame	32
			Local (Wait) Connection_Auth	None
			Local (Setup) Xmit_File	32
			Confirm	32

Name	Type	Command Sequence	ACK Msg. Size (bytes)
Receive_Source	Application	Local (Ansr) Ack_For_Session	32

Name	Type	Packets (bytes)	O.H. (bytes)	Retransmissions (msec.)
FXLP	Protocol	2016	32	500

Flow Control	Rate Control
None	None

NEXGENBUS LAB TESTBED

The lab set-up consists of two PC's, each of which is running the Microsoft NT 4.0 Service Pack 3 Operating System. The workstations are a Pentium II running at 333 MHz and a Pentium running at 200 MHz. They each have 64 Mbytes of memory and large hard disks. They have FibreXpress Systran PCI cards with FXLP drivers installed in a PCI card slot. There are twinax wires connecting the receive port on one card to the transmit port on the other card. There is a break out connector, which is used to add the Ancot Fibre channel protocol analyzer between the two PC's.

TEST PROCEDURES

The throughput and transfer times for different message sizes are measured in the lab with the WinBench program developed for NexGenBus. Each message size is transmitted from one PC and recorded on the receiving computer. The WinBench program is used to transmit 1000 messages of each message size. One thousand messages will be used to obtain a more precise number for the average transfer time and throughput. The message sizes were 32768, 65536, 196608, 1000000 bytes. These file sizes were originally chosen because they were easily coded in hex and are retained for consistency in the lab. Although larger messages could be transferred in the lab, it was found that to simulate file sizes larger than 1,000,000 bytes the simulations ran for an excessive amount of time.

The throughput and time to transfer are recorded for each file transfer. For each of the file transfers a fixed amount of delay is inserted between the file transfers. This delay allowed the processor of each PC to transfer the files to the buffer on the PCI card. Without this delay between file transfers the processor becomes the bottleneck in the system and prevents the datalink from being utilized to the maximum extent possible. The delays that are inserted between each different file transfer vary depending on when the throughput flattens out. In general, the delays inserted are 0, 1, 2, 3, 4, 6, 8, 10 and 20 milliseconds. Table 2 and 3 lists the file transfers sizes, recorded throughput, transfer time and time between messages.

The primary measurements to be used in the system analysis are throughput and latency. Throughput is defined as the amount of bandwidth allowed for a particular message. For example, if a 1000 byte message takes .001 seconds to go from one computer to another, the allowed bandwidth is reported as 1000 bytes/.001 or 1 Mbyte/sec (8 Mbits/sec). The latency is defined as the time for a message to be transported across the network from one computer to

another. The transfer times shown in Table 2 are the time when the file transfer is started on the source computer until the file is received on the target computer.

TEST RESULTS

Tables 2 and 3 show the results of the message transfer between the two PC's configured as shown in Figure 1. Column 1 is the size of the file transfer (number of Megabytes). Column 2 is the throughput (Megabytes/second) recorded in the file transfer. Column 3 is the time (seconds) to complete the file transfer and column 4 is the amount of delay (milliseconds) inserted between the individual file transfers. All file transfers were done 1000 times to achieve statistical averages. The 333 MHz PC is the source in the first set of data (Table 2) and the 200 MHz PC is the source in the second set of data (Table 3). The NexGenBus team is currently investigating how to better measure the latency with the WinBench program. The best throughput measurements with the delays added between the file transfers do not match between the 333 MHz and 200 MHz PC's. During the transfers, the CPU's are still responding to system interrupt calls. The faster PC can service those interrupts quicker thus spending more time on the file transfer.

Table 2
LabTest Results

Source PC: Pentium II 333 MHz, 64 Mbyte RAM

Msg Size (Mbytes)	Throughput(Megabytes/sec)	Xfer Time(sec)	Delay(ms)
.032768	21.31	1.57	0.0
.032768	21.16	1.56	1.0
.032768	45.00	2.00	2.0
.032768	45.17	3.00	3.0
.032768	44.87	4.00	4.0
.032768	44.90	7.00	7.0
.032768	44.72	10.12	10.0
.032768	45.02	20.00	20.0
.065536	23.85	2.76	0.0
.065536	23.73	2.76	1.0
.065536	23.72	2.78	2.0
.065536	47.11	3.01	3.0
.065356	66.55	4.00	4.0
.065536	66.39	5.00	5.0
.065536	67.23	6.00	6.0
.065536	61.87	10.06	10.0
.065536	66.71	20.00	20.0
.196608	29.29	6.73	0.0
.196608	29.33	6.72	2.0
.196608	29.31	6.73	4.0

.196608	29.26	6.75	6.0
.196608	73.00	8.00	8.0
.196608	73.75	10.01	10.0
.196608	73.63	20.00	20.0
1.000000	44.90	22.29	0.0
1.000000	44.90	22.29	2.0
1.000000	44.88	22.30	6.0
1.000000	44.88	22.30	10.0
1.000000	44.84	22.03	20.0
1.000000	67.84	30.02	30.0
1.000000	67.82	35.00	35.0
1.000000	67.82	40.00	40.0

Table 3

Source PC: Pentium 200 MHz, 64MbyteRAM

Msg Size (Mbytes)	Throughput(Megabytes/sec)	Xfer Time(sec)	Delay(ms)
.032768	18.16	1.85	0.0
.032768	18.14	1.85	1.0
.032768	18.56	2.05	2.0
.032768	19.04	3.01	3.0
.032768	19.23	4.00	4.0
.032768	19.54	7.00	7.0
.032768	20.51	10.00	10.0
.032768	23.53	19.98	20.0
.032768	26.61	29.90	30.0
.032768	30.00	39.96	40.0
.032768	29.10	49.95	50.0
.032768	30.90	59.94	60.0
.065536	23.73	2.80	0.0
.065536	23.73	2.80	2.0
.065536	28.11	4.00	4.0
.065536	28.22	6.00	6.0
.065536	28.50	8.00	8.0
.065536	27.89	10.01	10.0
.065536	28.24	19.99	20.0
.196608	30.39	6.51	0.0
.196608	30.59	6.46	4.0
.196608	30.48	6.49	6.0
.196608	42.60	8.00	8.0
.196608	42.89	10.00	10.0
.196608	42.50	19.90	20.0

1.000000	47.14	21.25	0.0
1.000000	47.18	21.24	2.0
1.000000	47.12	21.26	6.0
1.000000	47.13	21.26	10.0
1.000000	47.13	21.26	20.0
1.000000	52.18	29.90	30.0
1.000000	52.77	34.90	35.0
1.000000	52.79	39.90	40.0

Simulation Results

Table 4 shows the simulation results, which attempt to match the results of the lab using the 333 MHz PC. The first column lists the message size (Megabytes), the second column lists the throughput (Megabytes/sec), the third column is application latency (milliseconds), and the fourth column is transmission latency (milliseconds). The throughput results from the simulations match the best throughput measurements in the lab. This is a further indication that the Fibe channel datalink is not being fully utilized due to the PC's being restricted by their CPU's. The application latencies in the simulations match the application latencies in the lab file transfers at the lower file sizes (32768, 65536 and 196608). The 1 Mbyte file size lab transfers and the simulations show decreased throughput. This is a result of the 1 Mbyte buffer size in the Systran PCI card. Because the 1 Mbyte buffer can only hold one Mbyte file transfer at a time the throughput is restricted. The simulations and file transfer application latencies also diverge at the 1 Mbyte file size. The lab file transfers were between 30 and 40 milliseconds and the simulations predicted 10 milliseconds. The divergence on the application latencies is caused by the combination of the large delays inserted between the file transfers and the buffer restricting the throughput on the PCI card. Without the buffer causing the throughput to decrease, the throughput and latency would have correlated at the 1 Mbyte transfer. The transmission latencies (latency on the link only) predicted in the simulations (.013 msec.) match values predicted by the Systran card vendor. These values cannot be measured in the lab with the current configuration.

Table 4

Simulations of Point-to-Point File Transfers with 300 MHz PC

Msg Size (Mbytes)	Throughput(Mbytes/sec)	App. Latency (msec.)	Trans. Latency (msec.)
.032768	44.516160	2.55	.013
.065356	66.627552	5.10	.014
.196608	73.949937	15.30	.015
1.000000	66.599416	10.585	.015

TEST vs. SIMULATION ANALYSIS

Figure 2 is a graph of the throughput results from the simulation data (Table 4) and the lab test results (Table 2). The values used for throughput values are an average of when the throughput

has leveled off. The simulation and lab test data throughput measurements match within 5 % of each other. The fibre channel link during any of the file transfers file was never utilized over 50 %. The throughput measurement with no delay between the file transfers does not approach the theoretical predicted in the simulations. Figure 3 shows the latency measurements during the lab file transfer matches the application delays in the simulation up to the .196608 Megabyte file transfer. The one Megabyte file transfer has a longer delay because of the 1 Megabyte buffer size on the PCI card. The simulation matches the results of having the PC processor not restrict the throughput in which case the maximum throughput is realized and the only delay is the application latency and transmission delay on the link.

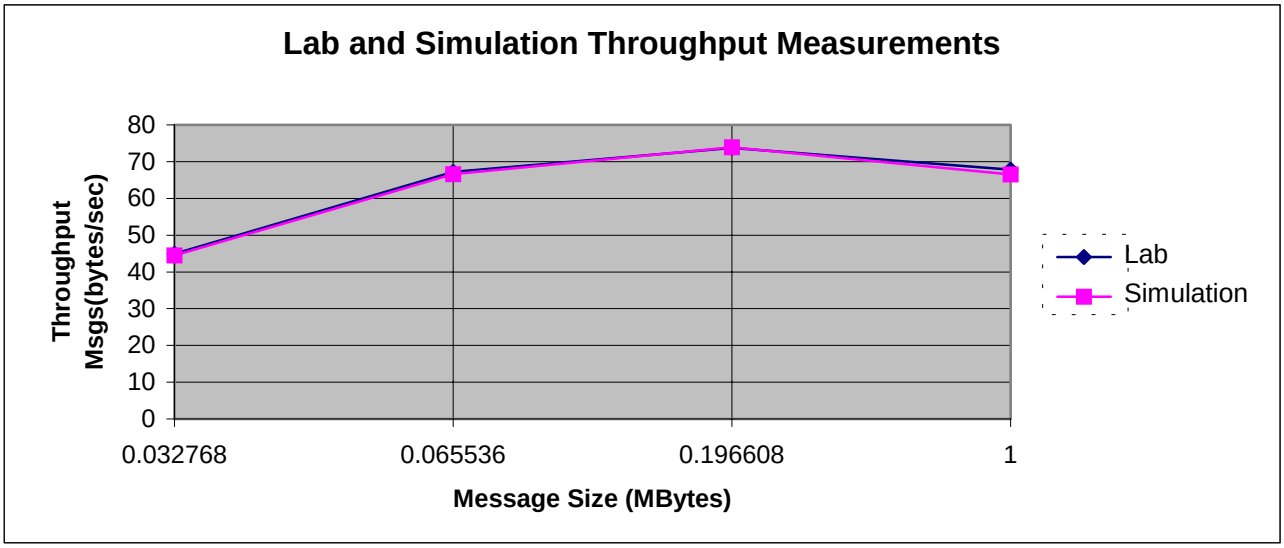

Figure 2

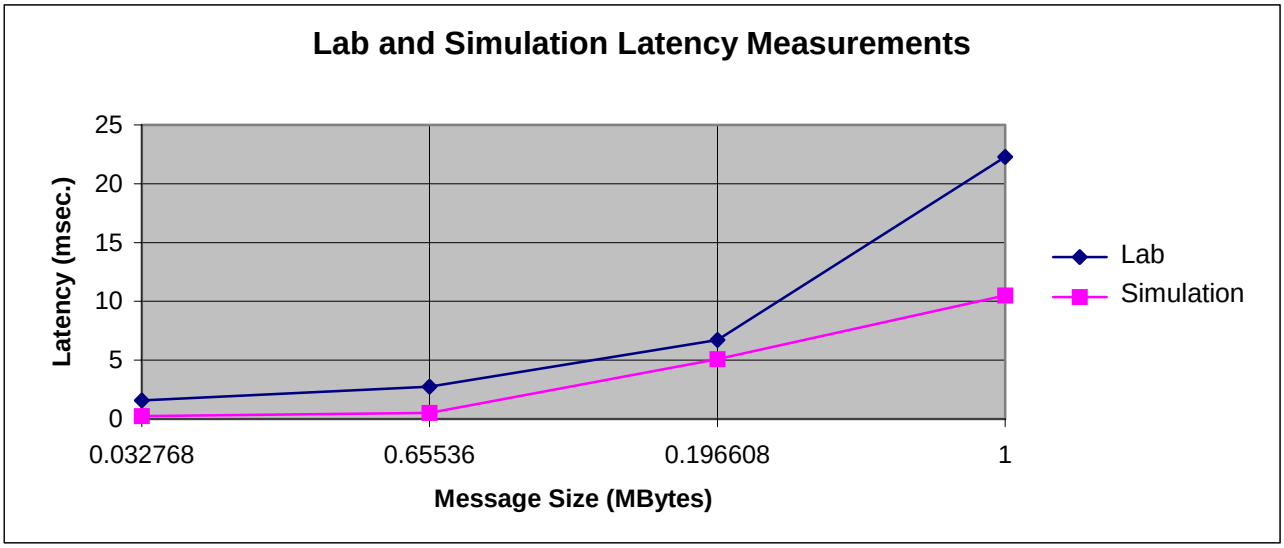

Figure 3

FOLLOW-ON EXPANDED CAPABILITY MODELS

The simulations have been verified with the arbitrated loop architecture in the lab. Because the lab setup was a simple two-node loop, many of the arbitrated loop functions were not present. This initial simulation model shows the results are consistent with the lab results. Therefore, the model can be expanded to include different classes of service, different protocols, additional nodes, synchronization and timing issues. As the model is expanded, additional lab verification may be required.